SECURIT

AD-A208 646

ENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unc	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT This document has been approved for public release and sale; its distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION U.S. Army Medical Research Institute of Infectious Diseases	6b. OFFICE SYMBOL (If applicable) SGRD-UIP-B
7a. NAME OF MONITORING ORGANIZATION U.S. Army Medical Research and Development Command	
6c. ADDRESS (City, State, and ZIP Code) Fort Detrick, Frederick, MD 21701-5011	7b. ADDRESS (City, State, and ZIP Code) Fort Detrick, Frederick, MD 21701-5012
8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

DTIC ELECTE JUN 07 1989 S D

11. TITLE (Include Security Classification)
Precursors of Androctonus australis Scorpion Neurotoxins

12. PERSONAL AUTHOR(S)
Drs. Pierre E. Bougis, Herve Rochat and Leonard A. Smith

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
	FROM _____ TO _____	24 MAY 89	31

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We isolated full-length cDNA clones encoding Androctonus australis scorpion neurotoxins active in mammals or insects. Sequence analysis of the cDNAs revelaed that precursors of toxins contained signal peptides of about 20 amino acids. In addition, the scorpion neurotoxins active on mammals had diverse peptide extensions at their C-terminal ends. Accordingly, the processing steps required for maturation into native toxins were not identical for all of the Androctonus australis neurotoxins. Southern blot analysis, performed at the genomic level with toxin II cDNA, seems to indicate that these toxin genes are unique. Finally, as the first successful attempt to express animal toxins, we demonstrated that monkey kidney COS-7 cells transfected with a toxin II cDNA clone transiently expressed a biologically active recombinant toxin.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

Precursors of Androctonus australis Scorpion Neurotoxins *

STRUCTURES OF PRECURSORS, PROCESSING OUTCOMES, AND EXPRESSION OF A FUNCTIONAL RECOMBINANT TOXIN II

Pierre E. Bougis, Herve Rochat[1] and Leonard A. Smith

from the Toxinology Department, Pathology Division, United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, MD 21701-5011, U.S.A. and [1]from the Universite d'Aix-Marseille II, Faculte de Medecine Secteur-Nord, Biochimie, CNRS UA1179 - INSERM U172, Bd. P. Dramard, 13326 Marseille Cedex 15, France.

*The research described was in compliance with the NIH guidelines involving recombinant DNA molecules. The investigators adhered to the "Guide for the Care and Use of Laboratory Animals," as promulgated by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animals Resources, National Research Council. The views of the authors do not purport to reflect the positions of the Department of the Army or the Department of Defense. During this work, Pierre E. Bougis held a National Research Council - USAMRIID Research Associateship from the National Academy of Sciences, and his permanent address for correspondence is: Universite d'Aux-Marseille II, Faculte de Medecine Secteur-Nord, Biochimie, CNRS UA1179 - INSERM U172, Bd. P. Dramard, 13326 Marseille Cedex 15, France.

We isolated full-length cDNA clones encoding Androctonus australis scorpion neurotoxins active in mammals or insects. Sequence analysis of the cDNAs revealed that precursors of toxins contained signal peptides of about 20 amino acids. In addition, the scorpion neurotoxins active on mammals had diverse peptide extensions at their C-terminal ends. Accordingly, the processing steps required for maturation into native toxins were not identical for all of the Androctonus australis neurotoxins. Southern blot analysis, performed at the genomic level with toxin II cDNA, seems to indicate that these toxin genes are unique. Finally, as the first successful attempt to express animal toxins, we demonstrated that monkey kidney COS-7 cells transfected with a toxin II cDNA clone transiently expressed a biologically active recombinant toxin.

The Buthid scorpion venoms are known to contain single-chain, basic proteins of 60 to 70 amino acid residues (1,2), tightly reticulated by four disulfide bridges (3). The high toxicity of these proteins is due to their high binding affinity for voltage-sensitive sodium channels, thus impairing the initial, rapid depolarization phase of the action potential in nerve and muscle. The pharmacology of these toxins has been reviewed elsewhere (4). Interestingly, these neurotoxins may be species-specific and the venom of a given species may contain toxins selectively active on mammals as well as others lethal to either insects or crustaceans (5,6).

During the last 30 years, numerous works have focused on the purification and characterization of immunochemical and structural relationships of diverse scorpion neurotoxins. Briefly, five distinct groups of toxins active in mammals have been defined according to both sequence and antigenic homologies (7). Considering their binding properties and pharmacology, they have been also classified into α- or ß-type toxins (8,9). Chemical modifications of assigned amino acid residues have allowed the determination of critical amino acid residues involved in the toxic activity (10). Four main antigenic epitopes appear responsible for immunogenicity and have been mapped by using homologous toxins (11), chemically modified toxins (12) and model synthetic peptides (13). Thus, while scorpion neurotoxins are polymorphic proteins, conserved domains exist which have been proposed to be the putative toxic site (14). The tridimensional structures of two scorpion neurotoxins have been determined (15,16).

For improved serotherapy and possible vaccine development, we wished to pursue the study of structure-activity relationships of scorpion

neurotoxins by carrying out specific amino acid substitutions in neurotoxin molecules by using recombinant DNA techniques. As a first step forward and considering also the lack of knowledge on the molecular biology of scorpion neurotoxins, we cloned and sequenced several cDNAs encoding diverse Androctonus australis scorpion neurtoxins. Based on the deduced primary structures and the directly determined amino acid sequences of the native toxins, the structures of toxin precursors suggested multiple post-translational processing steps. These steps were not identical for all toxins. The present paper describes experiments in which we transfected COS-7 cells with a construct containing an Androctonus australis toxin II cDNA in order to investigate the ability of COS-7 cells to express a biologically active recombinant toxin.

Keywords: toxin - antitoxin; (KT)

EXPERIMENTAL PROCEDURES

Materials- The following reagents were obtained from the indicated sources: [^{35}S] and [^{32}P]-labeled nucleotides, New England Nuclear, Boston, MA; ^{125}Iodide, carrier free, Amersham, Arlington Heights, IL; Restriction enzymes, M-MLV reverse transcriptase, terminal dideoxynucleotidyl transferase, Bethesda Research Laboratories, Gaithersburg, MD; commonly used molecular biology chemicals, Sigma, St. Louis, MO; Oligo d(T)-tailed pcDV-1-derived plasmid and oligo d(G)-tailed pL1 Hind III linker for cDNA cloning and expression, Escherichia coli, K12 MC1061 strain, Pharmacia, Piscataway, NJ; Sequenase, United States Biochemical Corporation, Cleveland, OH; X-AR and X-RP autoradiography films, Kodak, Rochester, NY.

Purification and Characterization of Telson mRNA-Scorpions of the species Androctonus australis were collected in Beni-Khedache, Tunisia, for the Pasteur Institute of Tunis. Animals were transported alive to USAMRIID where they were killed 2 days after manual extraction of their venom. The source of the mRNA was the telson, which is the last segment of the animal tail containing the two venom glands. Total RNA was extracted from homogenized telsons using the guanidinium hot-phenol method (17). Vanadyl-ribonucleoside-complex was used as an effective ribonuclease inhibitor. Poly (A)+ mRNA was further selected for by oligo (dT)-cellulose chromatography (18). The intactness of scorpion poly (A)$^+$ RNA and its ability to serve as template for full-length cDNA transcription was measured using 1 µg mRNA, 0.4 µg of oligo pd(T)$_{12-18}$ as primer and 200 units of M-MLV reverse-transcriptase. [^{32}P]-labeled primary transcripts were submitted to electrophoresis on a 1.5% alkaline agarose gel, blotted

onto Whatman DE-81 paper, and autoradiographed with X-AR film with intensifying screens for 48 h at -70 °C.

Construction of a cDNA Library-The Okayama-Berg cloning and expression system (19) was used to generate the library. From 9 µg of telson mRNA and 2 µg of oligo d (T)-tailed pcDV-1 plasmid primer, approximately 8 x 10^5 E. coli K12 MC1061 transformants were generated, which constituted the library, we used in this study.

Screening of the cDNA Library-For each screening experiment, 400,000 clones from the library were analyzed by oligonucleotide probes having sequence homology to regions coding for native scorpion toxins. Probes were synthesized on a Biosearch 8700 DNA synthesizer by beta-cyanoethyl phosphoramidite chemistry and were [^{32}P]-end-labeled by T4 polynucleotide kinase. High- and low-density screenings of bacterial colonies for recombinant plasmids were performed on nitrocellulose filters (20). Filters were prehybridized for 2 h at 37 °C in 6X SSC, pH 7.0, containing 1X Denhardts, 0.5% SDS, 100 µg/ml sheared and denaturated salmon sperm DNA, and 0.05 % sodium pyrophosphate (17). Filters were then hybridized in 6X SSC, pH 7.0, containing 1X Denhardts, 20 µg/ml yeast tRNA, 0.05 % sodium pyrophosphate, and the [^{32}P]-end-labeled oligonucleotide probe for 16 h at 37 °C. Successive washes were performed in 6X SSC, pH 7.0, containing 0.05 % sodium pyrophosphate at 37 °C for 1 hour and 47 °C for 5 min before autoradiography for 36 h at -70°C.

DNA Sequence Analysis-Inserts within BamH I and Pst I restriction sites adjacent to the poly A/T and poly G/C tracts of the pcD vector were

subcloned into both M13mp18 and M13mp19. Nucleotide sequence was determined by the dideoxy-sequencing method (21). Sequenase, ^{35}S-deoxyadenosine 5'-(α-thio) triphosphate, and the 17-base universal M13 primer were used.

Sequence analysis-Software from Intelligenetics, Inc. (Mountain View, CA), including SEQ, PEP, QUEST, and IFIND, were used on a VAX (Digital Equipment Corp., Lanham, MD).

DNA Transfection-Transfection of DNA was performed by the DEAE-dextran method using chloroquinine as described by Luthman and Magnusson (22). COS-7 (SV40-transformed African green monkey kidney) cells, obtained from American Type Culture Collection (CRL 1651, Passage Number: 9) were maintained at 37 °C under 5% CO_2 in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. Cells were plated onto 75-cm^2 plates and were subconfluent at the time of transfection. Cells were fed each day with 5 ml of Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum.

Southern Blot Analysis-High molecular weight genomic DNA was prepared from Androctonus australis scorpion muscle tissue. After complete digestion with appropriate restriction endonucleases, samples (10 μg) were electrophoresed in 0.8 % agarose gel. DNA was blotted onto nitrocellulose membrane and probed with BamH I/Pst I insert of pcD-402 plasmid encoding AaH II. The probe was [^{32}P]-labeled by the nick translation kit from Boehringer Mannheim. The filter was washed extensively in 0.1X SSC containing 0.1 % SDS.

Recombinant Toxin Immunopurification and Characterization-One ml of CNBr-activated Sepharose 4B from Pharmacia was used to prepare an

anti-AaH II immunoabsorbent column with 10 mg of antibodies purified from rabbit immunoserum (23) on a ImmunoPure column from Pierce. After elution with 0.1 N formic acid containing 0.15 M NaCl, the immunoabsorbed recombinant toxin was desalted on Sephadex G-10 in the presence of 0.1 N acetic acid containing 0.1% bovine serum albumin to minimize loss, then lyophilized. AaH II was purified as previously described (2) and it was iodinated with [^{125}I] by the lactoperoxidase method and purified by immunoprecipitation (24). AaH II-specific immunoassay was essentially as described elsewhere (25), except that 40 x 10^{-12} M of AaH II-specific antibodies and 6 x 10^{-12} M of [^{125}I]-AaH II were used. AaH II sensitive receptors were assayed on rat brain synaptosomes, as described (26).

In-Vivo Assay-Toxicity was assayed in male C57/BL_6 mice (Evic-Ceba, 33 Blanquefort, France) weighing 20 ± 3 grams by intracerebroventricular injection, as described by Haley and McCormick (27).

RESULTS

Isolation of cDNA Clones Encoding Toxins-The first step was to prepare the mRNA of scorpion telson from homogenized fresh telsons. Starting with 2 g of tissue (about 30 telsons), 42 µg of mRNA were obtained. In order to estimate the level of heterogeneity within the mRNA population, primary transcripts were synthesized by using oligo pd(T) primed reverse-transcriptase. After electrophoresis, numerous transcripts corresponding in size to about 360 bases and 1.100 bases were observed, the former being preponderant (Fig. 1). When scorpion venom is extracted before the venom gland is removed, the toxin-producing cells are stimulated to enter the secretory phase. Since we believe that 360 bases was the size of the mRNA encoding the 65-amino acid toxin, we did not proceed further to a size selection of telson mRNAs. Accordingly, in order to isolate cDNA clones encoding toxins, a large plasmid cDNA library was constructed by the procedure of Okayama/Berg in an expression vector containing a SV40 origin of replication (19)

The first toxin to be cloned successfully was AaH II. Two other screening experiments permitted the cloning of clone AaH I, AaH I', and AaH III, and AaH IT1 and AaH IT2. For each experiment, 400,000 transformants were plated and hybridized to oligonucleotide probe mixtures representing all possible complementary nucleotide sequences corresponding to the amino acid sequence of AaH II, AaH I or AaH IT (Table 1). Probes to AaH II were designed to maximize the amino acid sequence differences between AaH I and AaH II belonging, respectively, to the first and second structural groups. On the contrary, the probe to AaH I was designed to maximize

sequence homologies between all toxins belonging to the first structural group (AaH I, AaH I', AaH I", and AaH III). Using duplicate filters to screen the library, we obtained approximately 100 positve clones at each time. During the cloning process of AaH I, clones showing low-intensity hybridization were pooled separately from those of high-intensity. Thus, clones encoding AaH I' and AaH III were obtained. For each toxin cDNA being screened, clones were chosen at random, colony-purified, and analyzed for insert content. Figure 2 shows the physical map and size of the cDNA inserts from the clones analyzed. As indicated by the sizes of inserts, the percentage of full-length clones was high. Clone pcD-401 was shorter.

cDNA Nucleotide Sequences Encoding Toxins-The complete nucleotide sequence of inserts (within BamH I and Pst I sites) of 11 cDNA clones were determined by the dideoxy method with the strategy depicted in Fig. 2. Because of difficulties in sequencing through the poly A/T tract (except in the case of pcD-401) and because of limited enzyme restriction sites within cDNA sequences, we used synthetic oligonucleotides to extend the sequencing both ways.

The complete nucleotide sequences of representative cDNA clones encoding toxins active on mammals are shown in Fig. 3. The sequences, aligned for maximum homology with that of pcD-633, pcD-633 (AaH I), pcD-639 (AaH I'), and pcD-402 (AaH II), were probably full-length clones beginning with the same sequence, AACAA. pcD-635 and pcD-636 (Fig. 2) were identical to pcD-633. pcD-403 (Fig. 2) was 5' truncated (24 bases) and differed from pcD-402 by two substitutions: C/T (position 245),

corresponding to the third base of Cys codon, and T/C (position 371) in the 3' non-coding region. pcD-401 (Fig. 2) was also 5' truncated (starting at position 145) and was identical to pcD-403.

Fig. 4 shows the nucleotide sequences of representative cDNA clones encoding toxins active on insects. Clones pcD-644 (AaH IT1), pcD-648 (AaH IT1), as well as pcD-645 (AaH IT2) were almost full-length.

All these cDNA sequences contained one major open reading frame of about 310 nucleotides. At the 5' end, multiple potential initiation codons were observed. In each case, the 5' proximal ATG codon should have been responsible for the initiation of the translation (28). At the 3' end, a putative polyadenylation signal (AATAAA) was found 10 to 16 nucleotides upstream of the poly(A) tail (29). It appeared that 5' and 3' non-translated regions were approximately the same length.

Primary Structures of Toxin Precursors-For all clones, the open reading frame deomonstrated a coding capacity for toxin precursors larger in size than chromatographically characterized native toxins.

Precursors of toxins active on mammals possessed sequences extended at both N- and C- termini (Fig. 3). The proposed initiating Met was the first amino acid of a series of highly hydrophobic amino acids, suggesting the presence of a signal peptide of 19 amino acids. Homology among signal peptide sequences was very high. The signal peptide cleavage occurred for each precursor at Ser. Thus, the predicted N-terminal residue of each processed toxin was identical to tne one chemically determined on toxins isolated from the venom. An amino acid substitution was observed in the sequence of AaH III (at position 8) where Asp was found instead of the expected Asn. However, predicted C-terminal residues derived from the nucleotide sequences of cDNAs did not correlate with the C-terminal

residues obtained from direct amino acid sequence analysis. The toxin precursors ended with additional amino acids, respectively, Arg for AaH I, AaH I', and AaH III and Gly-Arg for AaH II.

Three distinct precursors of toxins active on insects were found (Fig. 4). For each of them, the signal peptide was 18 amino acids long and ended with Gly. The substitution Phe/Leu within the peptide signal (at position -2) was the only difference observed for the clones from which AaH IT1 was derived. A third clone contained a cDNA coding for the AaH IT2 toxin. This isoprotein differed from AaH IT1 by substitutions: Tyr/Asn (position 23) and Pro/Thr (position 67). In each case, Gln was found at position 25 instead of Glu. At the C-terminal ends, precursors had no additional amino acid residues.

Expression of Recombinant AaH II in Cos-7 Cells-The cloning vector was a transient eukaryotic expression system. We used the system to prove further that the pcD-403 clone was sufficient to direct the expression of a biologically active recombinant AaH II toxin. This plasmid was transfected into Cos-7 cells, the medium harvested between 24 and 168 h post-transfection, and assayed for the appearance of AaH II-related antigen by a specific radioimmunoassay (Fig. 5a). Medium from Cos-7 cells transfected with pcD-401 (a 5' truncated clone) had background values. The optimum quantity of plasmid used to transfect the Cos-7 cells was 5 µg and expression was maximum (0.2 µg/10^6 cells) at 120 h. These results demonstrated that pcD-403 was capable of directing expression of AaH II immunoreactive material in Cos-7 cells.

Immunological and Biological Properties of Recombinant AaH II-To purify the recombinant AaH II, the 120-h cell culture medium was chromato-

graphed on an immunoaffinity column. To ascertain whether or not the recombinant AaH II was as active in vivo as native AaH II, the affinity-purified product was injected intracerebroventricularly in mice. Mice died with symptoms clearly identical to those observed with mice dying from AaH II. From the results (Table 2), the concentration of the recombinant AaH II was estimated to be 6.9×10^{-8} M, with the assumption that its activity was the same as the one previously determined for AaH II (LD_{50}=0.5 ng/mouse) (30). The immunopurified recombinant AaH II was further characterized on the basis of its binding properties to both AaH II-specific antibodies (Fig. 5b) and rat brain voltage-sensitive sodium channels (Fig. 5c). The same serial dilutions of the sample were tested in parallel. Results established that the recombinant AaH II had the same ability as AaH II to compete in both assays with ^{125}I-AaH II. Thus, it is probable that the antigenic and toxic sites of the recombinant AaH II are closely related or identical to those of native AaH II.

Genomic Distribution of AaH II Sequence-To determine the size and the number of AaH II genes present in the scorpion genome, it was analyzed by Southern blot. High molecular weight DNA from scorpion muscle was digested with a variety of restriction enzymes, separated by agarose electrophoresis, transferred to nitrocellulose, and hybridized with a [^{32}P]-labeled pcD-402 insert. The results of this experiment are shown in Fig. 5. A single band was observed with EcoR I, Bgl II, Sph I, Sma I, Stu I, and Nru I restriction enzymes; the cDNA probe contained no restriction sites for these enzymes. Thus, the results were in favor of a unique copy gene for AaH II.

DISCUSSION

The range of Androctonus australis is North Africa. The venoms of animals collected in Chellala (Algeria) and Tozeur (Tunisia) have been extensively. The scorpion toxins comprise less than 2 % of the dried venom weight. Four toxins have been isolated from Androctonus australis which are active on mammals: AaH I and AaH I' (31), and AaH I" (only found in the venom of animals from Chellala) (32); and AaH II (33), AaH III (34), and AaH IV (35). Two toxins which are specifically toxic to insects have been isolated from the same species: AaH 1T1 (36) and AaH 1T2 (37). The amino acid sequences of AaH IV and AaH 1T2 have not been determined.

Size analysis of the mRNA pool obtained from the scorpion telson indicates that two main mRNA populations exist, and our initial assumption that 360 bases should be the expected size for monocystronic mRNA encoding toxin precursor has been experimentally verified. Indeed, by using synthetic oligonucleotides as screening probes, we isolated numerous full-length cDNA clones (about 370 base pairs) encoding these toxins including AaH IT2 but not AaH IV. Because the clones analyzed were picked randomly, they are representative of the toxin content of the venom, and the toxin precursors described here are the most probable and should not be considered as precursors of minor toxin isoforms.

Considering the cDNA-deduced amino acid sequences of toxin precursors, there exists as nearly a complete homology among signal peptide sequences within toxins active on mammals as with toxins active on insects. Signal peptide sequences are almost completely different when comparing precursors of these two types of toxins. These findings are

consistent with the classification of scorpion toxins, previously proposed at the level of their amino acid compositions, sequences and CD spectra, in which the toxin active on insects was distant from all the toxins active on mammals (38).

Sequence determination of toxin precursors gives an insight into the processing steps required to produce native toxins. Cleavage of the signal peptide by a signal-protease resulted in insecticidal toxins. For precursors of toxins active on mammals, two possibilities exist. For AaH II, in addition to the cleavage of the signal peptide, the removal of the C-terminal dipeptide Gly-Arg is required together with the α-amidation of the His (33). Precursors to secretory peptides synthesized as part of large and inactive precursor proteins contain sites for proteolysis and α-amidation. These are often marked by the sequence -X-Gly-B-B; X is the C-terminal amino acid in the mature peptide that is α-amidated, and B is either Lys or Arg (39). The additional dipeptide predicted by the nucleotide sequence encoding AaH II is in agreement with these observations as it has one basic residue instead of the usual two, or none at all, as in the case of melittin (40). The non- α-amidated scorpion toxins, active on mammals, not only require the removal of the signal peptide but also the removal of a C-terminal Arg residue, most probably by an exopeptidase. We have no explanation for the fact that different processing steps seem to exist for toxins that all are directed to the sodium channel.

Expression of recombinant AaH II was achieved by transfection of Cos-7 monkey kidney cells with an expression plasmid harboring a cDNA encoding AaH II. As far as we know, this is the first successful expression

of a recombinant animal toxin. The rationale for choosing a mammalian expression system was that, the biological activity of scorpion neurotoxins is, stricto sensu, dependent on the correct formation of disulfide bridges and the proper folding of the protein (41). Recombinant AaH II expressed by the Cos-7 cells was characterized in three different ways: by in-vivo assay, immunoassay, and receptorassay. The results obtained support the conclusion that Cos-7 cells expressed a recombinant protein that behaves as AaH II in each of the three assays. Therefore, we assume that the recombinant toxin that monkey kidney cells transiently expressed upon their transfection is the mature form of AaH II. The fact that the recombinant toxin was secreted in a biologically active form from the Cos-7 cells supports the scenario of signal peptide removal. Many bioactive peptides have carboxyl terminal α-amide residues and, in general, the presence of α-amide is critical for biological activity. We do not know what influence the α-amidation of the C-terminal His of AaH II has upon the biological activity of this toxin. Only a biochemical analysis of the recombinant toxin can certify its identity to AaH II and address the accuracy of the double processing of the toxin precursor. The level of expression obtained for the recombinant toxin did not permit such an analysis. We are, at present, attempting to increase the levels of expression of the recombinant toxin in other host systems to answer these questions.

Acknowledgements-The authors thank Drs Linda Lee and Catherine Loridan for their kind help, Marie-France Martin for her assistance with lethality assays, and Mohamed El Ayeb for providing us the scorpions.

REFERENCES

1. Miranda, F., and Lissitzky, S. (1961) Nature (London) **190**, 443
2. Miranda, F., Kopeyan, C., Rochat, H., Rochat, C., and Lissitzky, S. (1970) Eur. J. biochem. **16**, 514-523
3. Kopeyan, C., Martinez, G., Lissitzky, S., Miranda, F., and Rochat, H. (1974) Eur. J. Biochem. **47**, 483-489
4. Catterall, W. A. (1986) Annu. Rev. Biochem. **55**, 953-985
5. Zlotkin, E., Rochat, H., Kopeyan, C., Miranda, F., and Lissitzky S. (1971) Biochimie **53**, 1073-1078
6. Zlotkin, E., Miranda, F., and Lissitzky S. (1972) Toxicon **10**, 211-216
7. Rochat, H., Bernard, P., and Couraud, F. (1979) in Neurotoxins: Tools in Neurobiology, (Ceccarelli, B. and Clementi, F. eds.) Vol. 3, pp. 325-334 (Raven Press, New York).
8. Couraud, F., Jover, E., Dubois, J. M., and Rochat, H. (1982) Toxicon **20**, 9-16
9. Martin, M. F., Garcia y Perez, L. G., El Ayeb, M., Kopeyan, C., Bechis, G., Jover, E., and Rochat, H. (1987) J. Biol. Chem. **262**, 4452-4459
10. Darbon, H., Jover, E., Couraud, F., and Rochat, H. (1983) Int. J. Pept. Protein Res. **22**, 179-186
11. El Ayeb, M., Martin, M. F., Delori, P., Bechis, G., and Rochat, H. (1983) Mol. Immunol. **20**, 697-708
12. El Ayeb, M., Darbon, H., Bahraoui, E. M., Vargas, O., and Rochat, H. (1986) Eur. J. Biochem. **155**, 289-294
13. Bahraoui, E. M., Granier, C., Van Rietschoten, J., Rochat, H., and El Ayeb, M. (1986) J. Immuno. **136**, 3371-3377

14. Fontecilla-Camps, J. C., Almassy, R. J., Ealkick, S. E., Suddath, F. C., Watt, D. D., Feldman, R. J., and Bugg, C. E. (1981) Trends Biochem. Sci. **6**, 291-296

15. Fontecilla-Camps, J. C., Almassy, R. J., Suddath, F. L., Watt, D. D., and Bugg, C. E. (1980) Proc. Natl. Acad. Sci. USA **77**, 6496-6500

16. Fontecilla-Camps, J. C., Habersetzer-Rochat, C., and Rochat, H. (1988) Proc. Natl. Acad. Sci. USA **85**, 7443-7447

17. Maniatis, T., Fritsch, E. F., and Sambroock, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

18. Aviv, H. , and Leder, P. (1972) Proc. Nat. Acad. Sci. U.S.A. **69**, 1408-1412

19. Okayama, H., and Berg, P. (1983) Mol. Cell. Biol. **3**, 280-289

20. Hanahan, D. , and Meselson, M. (1980) Gene **10**, 63-67

21. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5467

22. Luthman, H., and Magnusson, G. (1983) Nucleic Acids Res. **11**, 1295-1308

23. Delori, P. Van Rietschoten, J. , and Rochat, H. (1981) Toxicon **19**, 393-407

24. Rochat, H., Tessier, M., Miranda, F., and Lisstzky, S. (1977) Anal. Biochem. **82**, 532-548

25. El Ayeb, M., Delori, P., and Rochat, H. (1983) Toxicon **21**, 709-716

26. Jover, E., Martin-Moutot, N., Couraud, F., and Rochat, H. (1978) Biochem. Biophys. Res. Commun. **85**, 377-382

27. Haley, T. J., and McCormick, W. G. (1957) Biochem. Biophys. Res. Commun. **95**, 1607-1614

28. Kozak, M. (1984) Nucl. Acids Res. **12**, 857-872

29. Proudfoot, N. J., and Brownlee, G. G. (1976) Nature **263**, 211-214

30. DeLima, M. E., Martin, M. F., Diniz, C. R., and Rochat, H. (1986) Biochem. Biophys. Res. Commun. **139**, 296-302

31. Rochat, H., Rochat, C., Kopeyan, C., Miranda, F., Lissitzky, S., and Edman, P. (1970) Eur. J. Biochem. **16**, 514-523

32. Martin, M. F., and Rochat, H. (1984) Toxicon **22**, 695-703

33. Rochat, H., Rochat, C., Sampieri, F., Miranda, F., and Lissitzky, S. (1972) Eur. J. Biochem. **28**, 381-388

34. Kopeyan, C., Martinez, G., and Rochat, H. (1979) Eur. J. Biochem. **94**, 609-615

35. Martin, M. F., and Rochat, H. (1986) Toxicon **24**, 1131-1139

36. Darbon, H., Zlotkin, E., Kopeyan, C., Van Rietschoten, J., and Rochat, H. (1982) Int. J. Pept. Protein Res. **20**, 320-330

37. De Dianous, S., Kopeyan, C., Bahraoui, E. M., and Rochat, H. (1987) Toxicon **25**, 731-741

38. Dufton, M. J., Drake, A. F., and Rochat, H. (1986) Biochim. Biophys. Acta **869**, 16-22

39. Mains, R. E., Eipper, B. A., Glembotski, C. C., and Dores, R. M. (1983), Trends Neuro. Sci., 229-235

40. Kreil, G., Mollay, C., Kaschnitz, R., Haiml, L., and Vilas, V. (1980) Ann. NY Acad. Sci. **343**, 338-345

41. Sabatier, J. M., Darbon, h., Fourquet, P., Rochat, H., and Van Rietschoten, J. (1987) Int. J. Peptide Protein Res. **30**, 125-134

Table 1. Synthetic oligonucleotides used for hybridization probes. All possible amino acid sequences for Androctonus australis toxins and sequences of hybridization probes are shown. I is inosine.

```
          8                                                    19
AaH I     Tyr Pro Asn Asn Cys Val Tyr His Cys Val Pro Pro

       5' TAC CCI AAC AAC TGT GTI TAC CAC TGT GTI CCI CC 3'
            :      T   T          T   T

AaH I"    --- --- --- --- --- --- --- --- --- --- --- ---

AaH I'    --- --- --- --- --- --- --- --- --- Ile --- ---

AaH III   Asn Ser Lys --- --- --- --- --- --- --- --- ---
```

```
          21                                   30                                   39
AaH II    Tyr Cys Asn Glu Glu Cys Thr Lys Leu Lys Gly Glu Ser Gly Tyr Cys Gln Trp Ala

       5' TAC TGC AAC GAA GAA TGC AC 3'                    5' GGA TAC TGC CAA TGG GC 3'
            T   T   T   G   G   T                              C   T   T   G
                                                               G
                                                               T

          24                        30                   35
AaH IT    Asn Gln Cys Thr Lys Val His Tyr Ala Asp Lys Gly

       5' AAC CAA TGC ACI AAA GTI CAC TAT GCI GAC AAA GG 3'
            T   G   T              T          T
```

Table 2. *In Vivo* activity of recombinant AaH II. Samples (5 µl) were injected intracerebroventricularly to mice. According to a LD_{50} of 0.5 ng / mouse for AaH II (30) the concentration of the recombinant toxin was estimated to be $6.9\ 10^{-8}$ M.

Sample dilution	Dead / Injected
none	2/2
1/5	3/5
1/10	0/3

FIGURE LEGENDS

Fig. 1. **Characterization by size of primary transcripts of the polyadenylated mRNA population purified from the venom glands of the scorpion Androctonus australis.** Given in basepairs, the size markers consist of 1 Kb DNA ladder from Bethesda Research Laboratory.

Fig. 2. **Sequencing strategy for cDNA clones encoding toxins.** A scale is given in base pairs (bp). The 5' and 3' ends are indicated. Direction and length of individual sequencing reactions are shown by horizontal arrows under each clone. The following primers were synthesized and used to prime M13 clones as indicated respectively:

(1) 5'GCGACGGTTTATGTAAG3',
(2) 5'CAAAGGATATTGCTGCT3',
(3) 5'AGTTGAAAGGTGAGAGT3',
(4) 5'TACGTACATGATCGGGC3',
(5) 5'TCGGTACGTTATCGGGC3',
(6) 5'GTCATTTAGACCGAAGC3'.

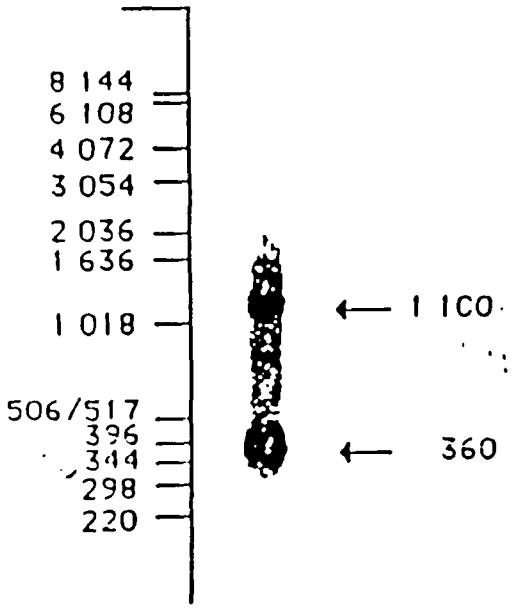
Fig. 3. **Nucleotide and predicted amino acid sequence of cDNA clones for mammal-specific toxins.** The nucleotide sequences beginning with the 5' end of the cDNA inserts are represented in the 5' to 3' direction and numbered on the top. Sequences are aligned for maximum homology with that of pcD-633 and differing nucleotides in each sequence are indicated on the top by symbols. The predicted protein sequences are given below the nucleotide sequences and are numbered starting at the

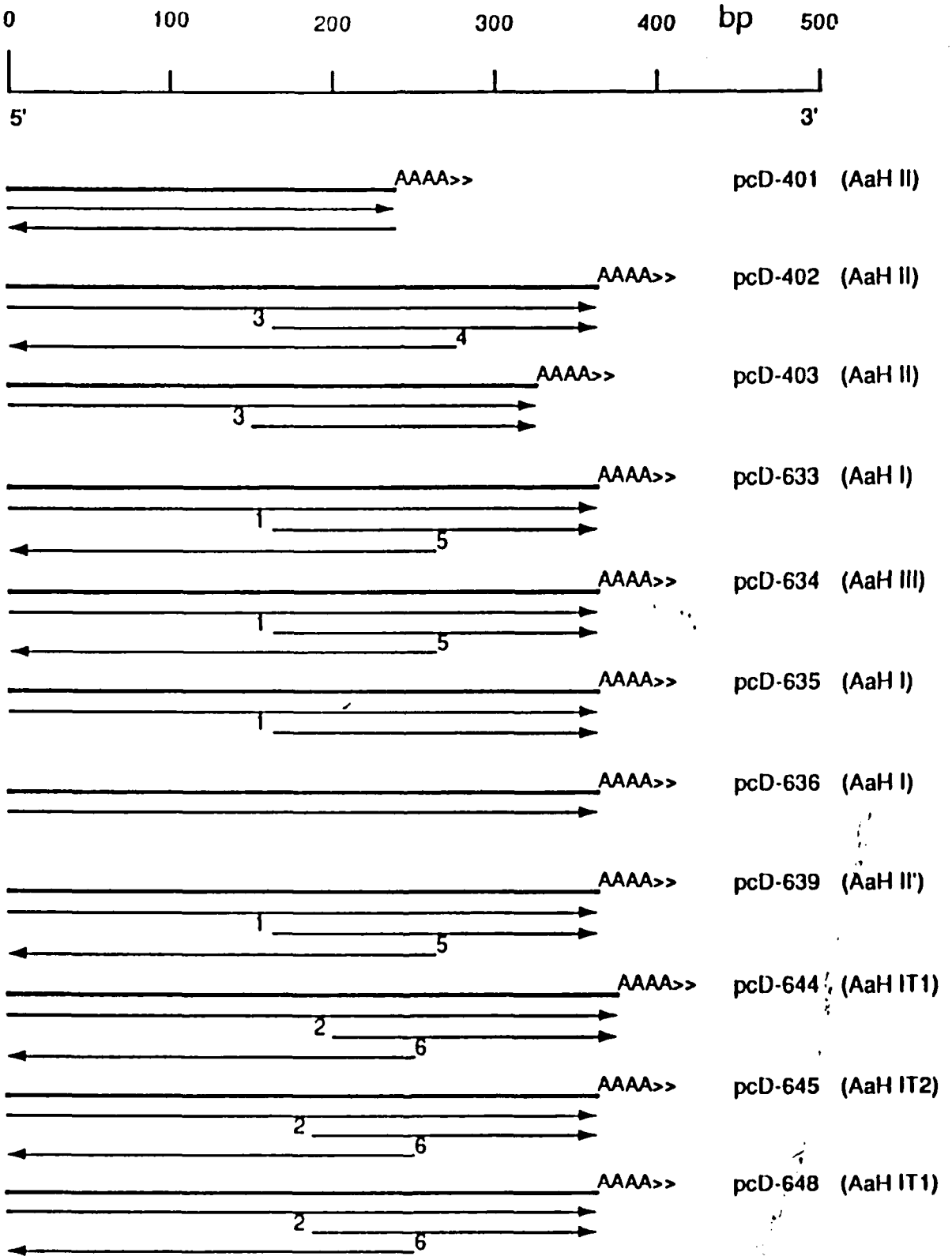
first amino acid of the mature protein. Signal peptide sequences are underlined. A potential polyadenylation signal of AATAAA is italicized.

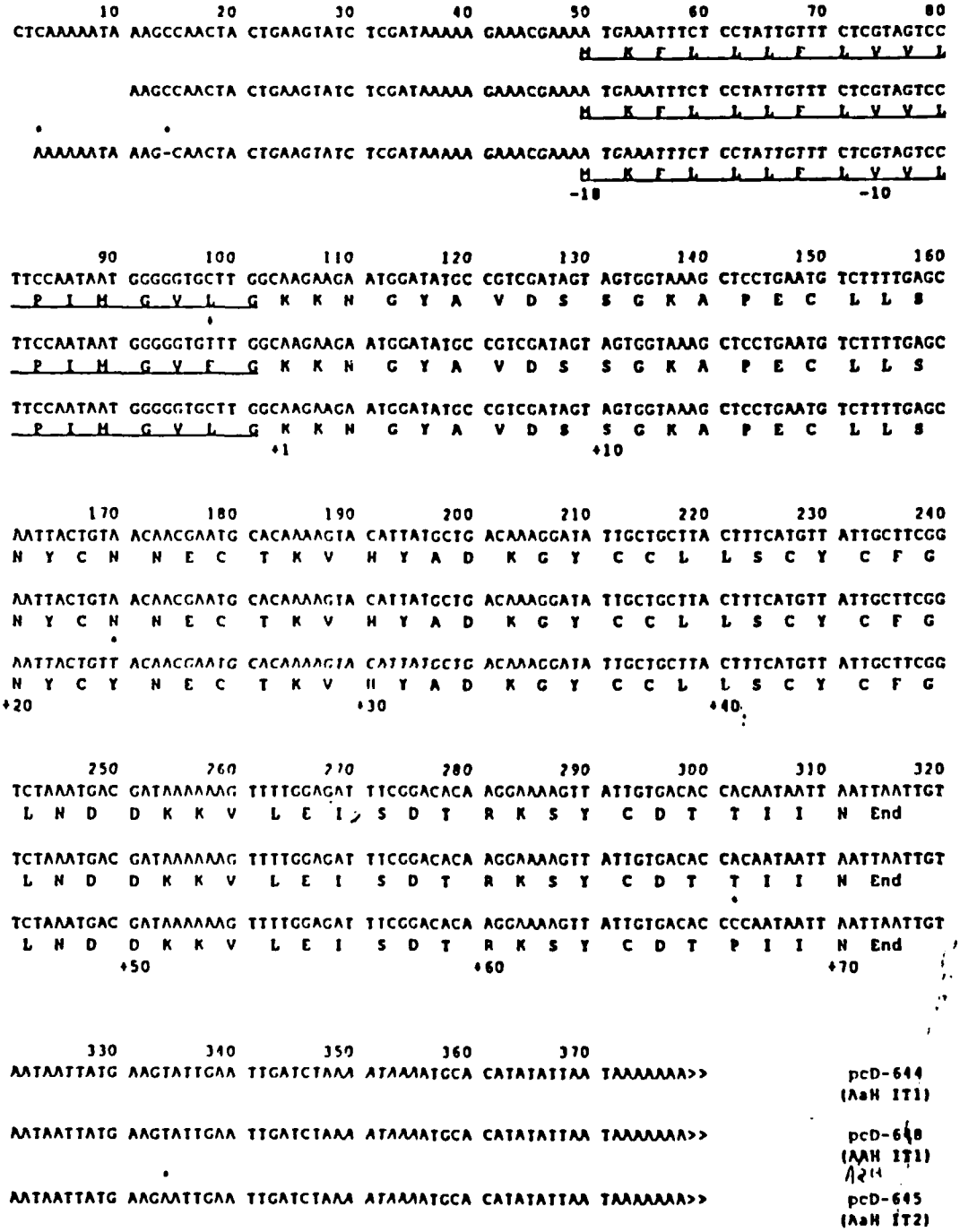
Fig. 4. **Nucleotide and predicted amino acid sequence of cDNA clones for insect-specific toxins.** The nucleotide sequences beginning with the 5' end of the cDNA inserts are represented in the 5' to 3' direction and numbered on the top. Sequences are aligned for maximum homology with that of pcD-644 and differing nucleotides in each sequences are indicated on the top by symbols. The predicted protein sequences are given below the nucleotide sequences and are numbered starting at the first amino acid of the mature protein. Signal peptide sequences are underlined. A potential polyadenylation signal of AATAAA is in italicized.

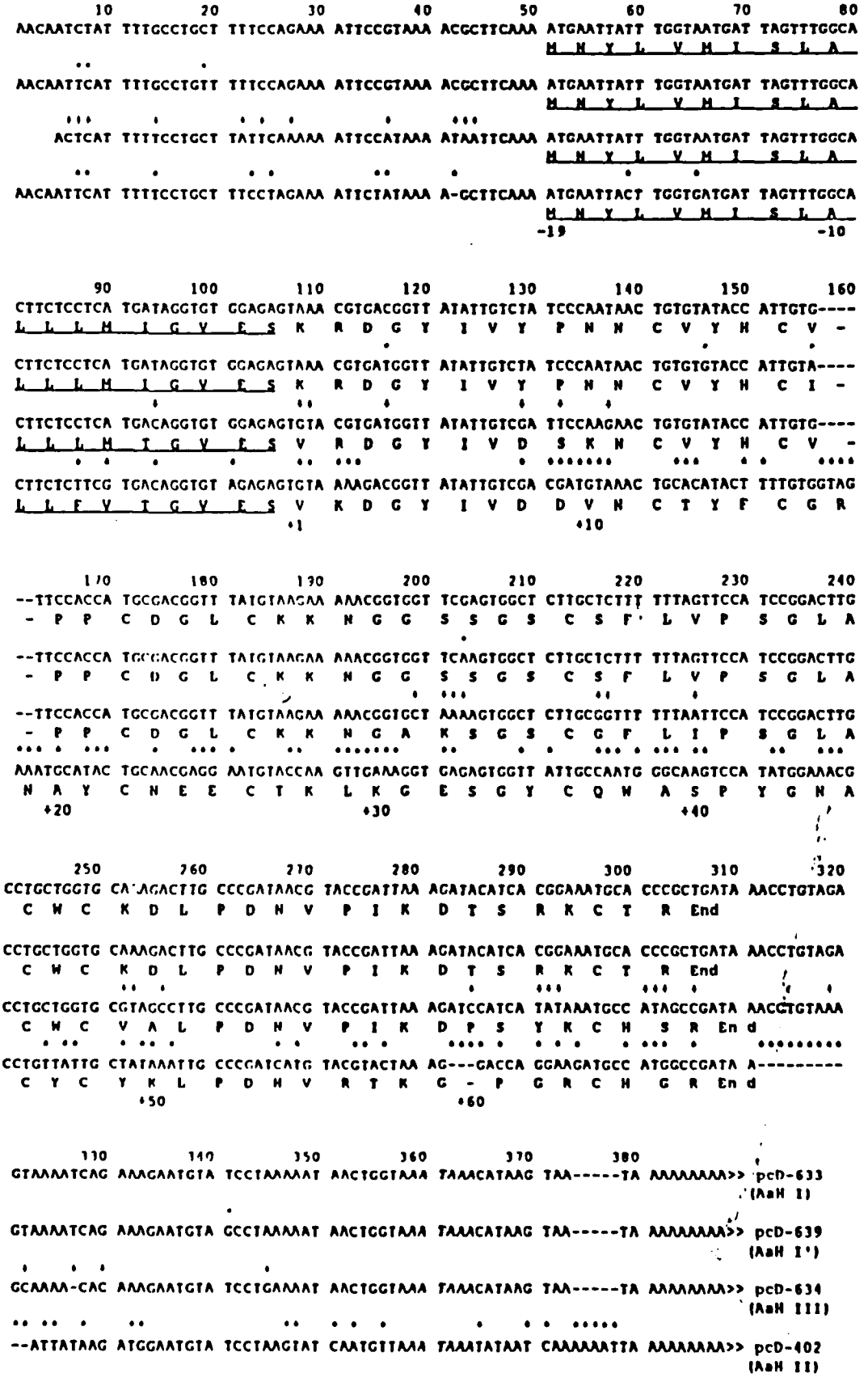
Fig. 5. **Expression in Cos-7 cell and immunological and biological characterization of recombinant AaH II.** (a) Monkey kidney cells were transfected with (O) 2 µg, (˙) 5 µg, or (Δ) 10 µg of recombinant plasmid pcD-403 or (◊) no-DNA; the expression of the recombinant toxin was followed on a daily basis by AaH II-specific immunoassays of cell culture medium. After being retrieved from cell culture media by immunoaffinity chromatography, the recombinant toxin was further characterized by (b) AaH II-specific immunoassays and (c) receptor assays. For both assays, (O) depicts the standard curve. Ten-fold serial dilutions of the immunopurified sample of recombinant AaH II (initial concentration of 6.9×10^{-8} M as estimated from the in-vivo experiment of Table 2) were in both assays and (X) are the results of such experiments.

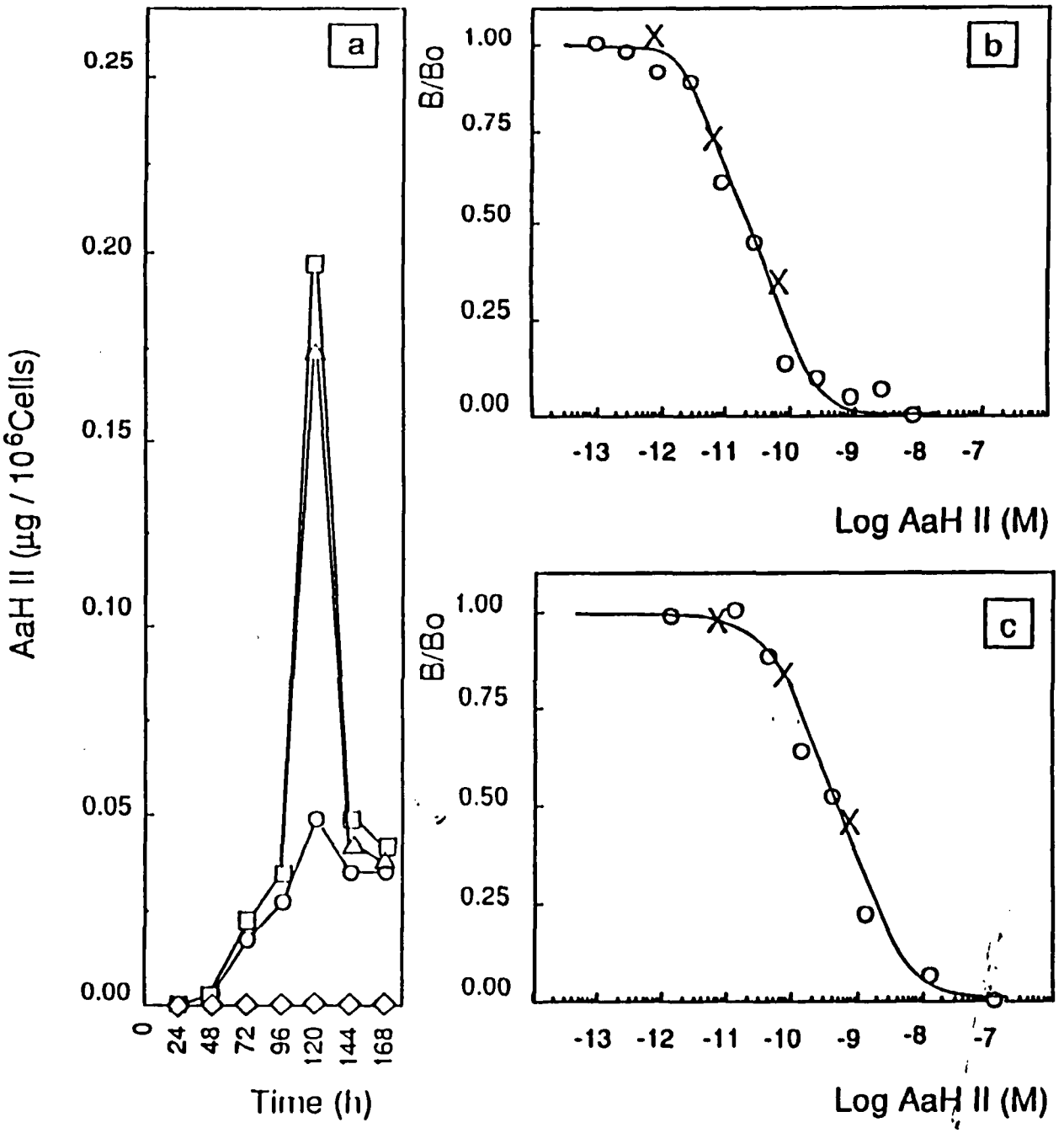
Fig. 6. **Southern blot analysis of Androctonus australis DNA.** Muscle DNA samples (10 µg) were digested with EcoR I, Bgl I, Sph I, Sma I, Xho I, Stu I, and Nru I and then electrophoresed on an 0.8 % agarose gel. The digests were then blotted on nitrocellulose filter and probed with BamH I/Pst I insert of pcD-402 plasmid encoding AaH II. The size markers, given in basepairs, consist of a 1 Kb DNA ladder from Bethesda Research Laboratory.


8 144
6 108
4 072
3 054
2 036
1 636
1 018
506/517
396
344
298
220
1 1CO
360


0
100
200
300
400
bp
500
5'
3'
AAAA>>
pcD-401 (AaH II)
AAAA>>
pcD-402 (AaH II)
3
4
AAAA>>
pcD-403 (AaH II)
3
AAAA>>
pcD-633 (AaH I)
1
5
AAAA>>
pcD-634 (AaH III)
1
5
AAAA>>
pcD-635 (AaH I)
1
AAAA>>
pcD-636 (AaH I)
AAAA>>
pcD-639 (AaH II')
1
5
AAAA>>
pcD-644 (AaH IT1)
2
6
AAAA>>
pcD-645 (AaH IT2)
2
6
AAAA>>
pcD-648 (AaH IT1)
2
6


10 20 30 40 50 60 70 80
CTCAAAAATA AAGCCAACTA CTGAAGTATC TCGATAAAAA GAAACGAAAA TGAAATTTCT CCTATTGTTT CTCGTAGTCC
M K F L L L F L V V L
AAGCCAACTA CTGAAGTATC TCGATAAAAA GAAACGAAAA TGAAATTTCT CCTATTGTTT CTCGTAGTCC
M K F L L L F L V V L
AAAAAATA AAG-CAACTA CTGAAGTATC TCGATAAAAA GAAACGAAAA TGAAATTTCT CCTATTGTTT CTCGTAGTCC
M K F L L L F L V V L
-18 -10
90 100 110 120 130 140 150 160
TTCCAATAAT GGGGGTGCTT GGCAAGAAGA ATGGATATGC CGTCGATAGT AGTGGTAAAG CTCCTGAATG TCTTTTGAGC
P I M G V L G K K N G Y A V D S S G K A P E C L L S
TTCCAATAAT GGGGGTGTTT GGCAAGAAGA ATGGATATGC CGTCGATAGT AGTGGTAAAG CTCCTGAATG TCTTTTGAGC
P I M G V F G K K N G Y A V D S S G K A P E C L L S
TTCCAATAAT GGGGGTGCTT GGCAAGAAGA ATGGATATGC CGTCGATAGT AGTGGTAAAG CTCCTGAATG TCTTTTGAGC
P I M G V L G K K N G Y A V D S S G K A P E C L L S
+1 +10
170 180 190 200 210 220 230 240
AATTACTGTA ACAACGAATG CACAAAAGTA CATTATGCTG ACAAAGGATA TTGCTGCTTA CTTTCATGTT ATTGCTTCGG
N Y C N N E C T K V H Y A D K G Y C C L L S C Y C F G
AATTACTGTA ACAACGAATG CACAAAAGTA CATTATGCTG ACAAAGGATA TTGCTGCTTA CTTTCATGTT ATTGCTTCGG
N Y C N N E C T K V H Y A D K G Y C C L L S C Y C F G
AATTACTGTT ACAACGAATG CACAAAAGTA CATTATGCTG ACAAAGGATA TTGCTGCTTA CTTTCATGTT ATTGCTTCGG
N Y C Y N E C T K V H Y A D K G Y C C L L S C Y C F G
+20 +30 +40
250 260 270 280 290 300 310 320
TCTAAATGAC GATAAAAAAG TTTTGGAGAT TTCGGACACA AGGAAAAGTT ATTGTGACAC CACAATAATT AATTAATTGT
L N D D K K V L E I S D T R K S Y C D T T I I N End
TCTAAATGAC GATAAAAAAG TTTTGGAGAT TTCGGACACA AGGAAAAGTT ATTGTGACAC CACAATAATT AATTAATTGT
L N D D K K V L E I S D T R K S Y C D T T I I N End
TCTAAATGAC GATAAAAAAG TTTTGGAGAT TTCGGACACA AGGAAAAGTT ATTGTGACAC CCCAATAATT AATTAATTGT
L N D D K K V L E I S D T R K S Y C D T P I I N End
+50 +60 +70
330 340 350 360 370
AATAATTATG AAGTATTGAA TTGATCTAAA ATAAAATGCA CATATATTAA TAAAAAAA>> pcD-644 (AaH IT1)
AATAATTATG AAGTATTGAA TTGATCTAAA ATAAAATGCA CATATATTAA TAAAAAAA>> pcD-648 (AAH IT1)
AATAATTATG AAGAATTGAA TTGATCTAAA ATAAAATGCA CATATATTAA TAAAAAAA>> pcD-645 (AaH IT2)


10 20 30 40 50 60 70 80
AACAATCTAT TTTGCCTGCT TTTCCAGAAA ATTCCGTAAA ACGCTTCAAA ATGAATTATT TGGTAATGAT TAGTTTGGCA
M N Y L V M I S L A
AACAATTCAT TTTGCCTGTT TTTCCAGAAA ATTCCGTAAA ACGCTTCAAA ATGAATTATT TGGTAATGAT TAGTTTGGCA
M N Y L V M I S L A
ACTCAT TTTTCCTGCT TATTCAAAAA ATTCCATAAA ATAATTCAAA ATGAATTATT TGGTAATGAT TAGTTTGGCA
M N Y L V M I S L A
AACAATTCAT TTTTCCTGCT TTCCTAGAAA ATTCTATAAA A-GCTTCAAA ATGAATTACT TGGTGATGAT TAGTTTGGCA
M N Y L V M I S L A
-19 -10
90 100 110 120 130 140 150 160
CTTCTCCTCA TGATAGGTGT GGAGAGTAAA CGTGACGGTT ATATTGTCTA TCCCAATAAC TGTGTATACC ATTGTG----
L L L M I G V E S K R D G Y I V Y P N N C V Y H C V -
CTTCTCCTCA TGATAGGTGT GGAGAGTAAA CGTGATGGTT ATATTGTCTA TCCCAATAAC TGTGTGTACC ATTGTA----
L L L M I G V E S K R D G Y I V Y P N N C V Y H C I -
CTTCTCCTCA TGACAGGTGT GGAGAGTGTA CGTGATGGTT ATATTGTCGA TTCCAAGAAC TGTGTATACC ATTGTG----
L L L M T G V E S V R D G Y I V D S K N C V Y H C V -
CTTCTCTTCG TGACAGGTGT AGAGAGTGTA AAAGACGGTT ATATTGTCGA CGATGTAAAC TGCACATACT TTTGTGGTAG
L L F V T G V E S V K D G Y I V D D V N C T Y F C G R
+1 +10
170 180 190 200 210 220 230 240
--TTCCACCA TGCGACGGTT TATGTAAGAA AAACGGTGGT TCGAGTGGCT CTTGCTCTTT TTTAGTTCCA TCCGGACTTG
- P P C D G L C K K N G G S S G S C S F L V P S G L A
--TTCCACCA TGCGACGGTT TATGTAAGAA AAACGGTGGT TCAAGTGGCT CTTGCTCTTT TTTAGTTCCA TCCGGACTTG
- P P C D G L C K K N G G S S G S C S F L V P S G L A
--TTCCACCA TGCGACGGTT TATGTAAGAA AAACGGTGCT AAAAGTGGCT CTTGCGGTTT TTTAATTCCA TCCGGACTTG
- P P C D G L C K K N G A K S G S C G F L I P S G L A
AAATGCATAC TGCAACGAGG AATGTACCAA GTTGAAAGGT GAGAGTGGTT ATTGCCAATG GGCAAGTCCA TATGGAAACG
N A Y C N E E C T K L K G E S G Y C Q W A S P Y G N A
+20 +30 +40
250 260 270 280 290 300 310 320
CCTGCTGGTG CAAAGACTTG CCCGATAACG TACCGATTAA AGATACATCA CGGAAATGCA CCCGCTGATA AACCTGTAGA
C W C K D L P D N V P I K D T S R K C T R End
CCTGCTGGTG CAAAGACTTG CCCGATAACG TACCGATTAA AGATACATCA CGGAAATGCA CCCGCTGATA AACCTGTAGA
C W C K D L P D N V P I K D T S R K C T R End
CCTGCTGGTG CGTAGCCTTG CCCGATAACG TACCGATTAA AGATCCATCA TATAAATGCC ATAGCCGATA AACCTGTAAA
C W C V A L P D N V P I K D P S Y K C H S R End
CCTGTTATTG CTATAAATTG CCCGATCATG TACGTACTAA AG---GACCA GGAAGATGCC ATGGCCGATA A---------
C Y C Y K L P D H V R T K G - P G R C H G R End
+50 +60
330 340 350 360 370 380
GTAAAATCAG AAAGAATGTA TCCTAAAAAT AACTGGTAAA TAAACATAAG TAA-----TA AAAAAAAA>> pcD-633
(AaH I)
GTAAAATCAG AAAGAATGTA GCCTAAAAAT AACTGGTAAA TAAACATAAG TAA-----TA AAAAAAAA>> pcD-639
(AaH I')
GCAAAA-CAC AAAGAATGTA TCCTGAAAAT AACTGGTAAA TAAACATAAG TAA-----TA AAAAAAAA>> pcD-634
(AaH III)
--ATTATAAG ATGGAATGTA TCCTAAGTAT CAATGTTAAA TAAATATAAT CAAAAAATTA AAAAAAAA>> pcD-402
(AaH II)


a
AaH II (µg / 10⁶Cells)
0.25
0.20
0.15
0.10
0.05
0.00
0
24
48
72
96
120
144
168
Time (h)
b
B/Bo
1.00
0.75
0.50
0.25
0.00
-13
-12
-11
-10
-9
-8
-7
Log AaH II (M)
c
B/Bo
1.00
0.75
0.50
0.25
0.00
-13
-12
-11
-10
-9
-8
-7
Log AaH II (M)

EcoR I
Bgl II
Sph I
Sma I
Xho I
Stu I
Nru I
12 216
6 108
5 090
4 072
3 054
2 036
1 636
1 018